Health Monitoring Survey of Bell 412EP Transmissions

Brian E. Tucker
Associate Technical Fellow
Bell Helicopter
Fort Worth, TX, USA

Paula J. Dempsey
Aerospace Research Engineer
NASA Glenn Research Center
Cleveland, OH, USA

ABSTRACT

Health and usage monitoring systems (HUMS) use vibration-based Condition Indicators (CI) to assess the health of helicopter powertrain components. A fault is detected when a CI exceeds its threshold value. The effectiveness of fault detection can be judged on the basis of assessing the condition of actual components from fleet aircraft. The Bell 412 HUMS-equipped helicopter is chosen for such an evaluation. A sample of 20 aircraft included 12 aircraft with confirmed transmission and gearbox faults (detected by CIs) and eight aircraft with no known faults. The associated CI data is classified into "healthy" and "faulted" populations based on actual condition and these populations are compared against their CI thresholds to quantify the probability of false alarm and the probability of missed detection. Receiver Operator Characteristic analysis is used to optimize thresholds. Based on the results of the analysis, shortcomings in the classification method are identified for slow-moving CI trends. Recommendations for improving classification using time-dependent receiver-operator characteristic methods are put forth. Finally, lessons learned regarding OEM-operator communication are presented.

INTRODUCTION

The integrity of a helicopter's powertrain is vital to the helicopter's mission and safety. Today's helicopters have been equipped with HUMS (Health and Usage Monitoring Systems) to detect damage in dynamic components. These systems consist of a main computer for data processing as well as sensors for the acquisition of in-flight data as shown in Figure 1. Sensors are typically installed on the airframe, gear boxes, and drive train components to identify any changes in vibration patterns over the course of flight operation [1]. The health of components is determined from vibration signatures that form the basis of component Condition Indicators (CIs). The CIs identify the presence of fault patterns that correspond to specific damage types on a component [2].

Condition Indicators

Dozens of CIs are used in HUMS. Table 1 lists example CIs and cites references that detail the associated algorithms. Shafts and gears produce vibration signals that repeat with each revolution of the component. Time synchronous averaging, or cycle averaging, divides the vibration signal into single revolution segments to produce a single "average" cycle. The result emphasizes periodic vibration trends and averages out noise [3]. Many shaft and gear CIs use time synchronous average data. The DA1 CI, in Table 1, calculates the root mean square (RMS) of the time synchronous average signal [4]. Other CIs convert the time synchronous average data into the frequency domain and extract signal amplitudes at particular frequencies of interest. These frequencies may be shaft speed (1/rev), gear mesh (GM) and its first harmonic (GM x 2), or sidebands adjacent to gear mesh tones (SI) [5,6]. Bearing vibration CIs often use asynchronous vibration data. Another example is the "High Frequency" bearing CI, which is the RMS of high-pass filtered asynchronous data [7].

HUMS calculate many CIs for each monitored component. The HUMS then uses CI thresholds to diagnose faults. Thresholds are based on past experience and engineering judgment. As shown in Figure 2, when a CI value exceeds the pre-defined threshold, damage is likely to be present and the HUMS will alert ground personnel to the detected fault.

Table 1. Example Condition Indicators

Name	Description	References
DA1	Diagnostic Algorithm 1 (RMS)	[4]
1/Rev	Amplitude at Shaft 1/revolution frequency	[5]
GM	Amplitude at Gear Mesh Frequency	[5]
2 x GM	Amplitude at 2 x Gear Mesh Frequency	[5]
SI	Sideband Index	[6]
HighFreq	"High Frequency" Bearing CI	[5],[7]

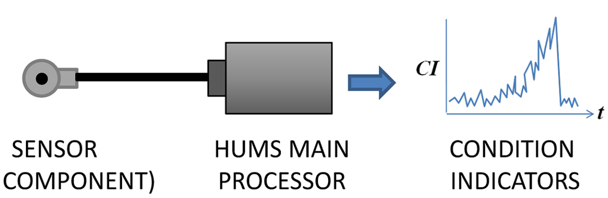


Figure 1. HUMS Generation of CI Values

Presented at the AHS 72nd Annual Forum, West Palm Beach, Florida, USA, May 17-19, 2016.

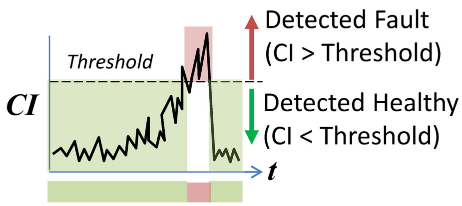


Figure 2. Detected Component Condition over Time Based on CI

Measuring CI Effectiveness

The effectiveness of CIs in detecting damage must ultimately be judged by results from fleet operations. "Truth data" comes from tear down analysis (TDA) of removed components. As shown in Figure 3, a CI is effective is when its detected condition matches the TDA actual condition (either a true healthy detection or a true fault detection). A CI is ineffective when its detected condition contradicts the actual condition. This "ineffectivity" takes two forms—either the CI detects a fault that cannot be confirmed by TDA (false alarm), or it does not detect a fault, yet a fault is confirmed by TDA (missed detection). These effectiveness measures may drive changes in CI thresholds or CI algorithms to improve HUMS performance. Reducing a CI threshold will reduce missed detections, but doing so may also increase false alarms. Previous investigations assessed drive train CI performance based on faulted components in military aircraft [6,8,9,10,11].

SCOPE

The current study applies CI performance analysis to commercial aircraft HUMS data. Many commercial operators use HUMS data as the basis for taking maintenance actions [12]. CI algorithms and thresholds can therefore strongly impact maintenance cost, especially when operating rules and industry best practices [13] require inspection or removal of drive system components based on HUMS data alone (in an effort to enhance safety).

For over 10 years, Health & Usage Monitoring Systems (HUMS) have been fielded on Bell 412EP aircraft [14]. The 412EP HUMS system is illustrated in Figure 4, with sensor locations shown around the airframe. A Honeywell 1209 Modern Signal Processing Unit (MSPU) is used for data acquisition and processing of sensor data. This is the same unit used in previous military aircraft HUMS performance studies. With nearly 150 HUMS-equipped 412EP aircraft in service, a good opportunity exists to explore CI effectiveness in detecting faulted components on a commercial aircraft. While many drive train components are monitored by the HUMS, only the CIs related to transmissions/gearboxes are discussed in this paper.

APPROACH

To restate the objective of this exercise, an evaluation of CIs from 20 HUMS equipped commercial Bell 412EP helicopters is being performed. The goal is to determine the efficacy of the CIs and the associated CI thresholds in detecting a fault (as confirmed by a Tear Down Assessment of the identified faulty component). The approach used to assess the CI performance includes:

- Identifying the CI(s) that triggered gearbox inspection or removal and gathering this CI data.
- Determining the times during which the gearbox was actually healthy or faulted (based on TDA findings) and to categorize the CI data into healthy and faulted CI data populations based on those time periods.
- Assessing how often the CI detected condition matched the actual condition (actually healthy or actually faulted) through statistical analysis of the healthy and faulted populations.

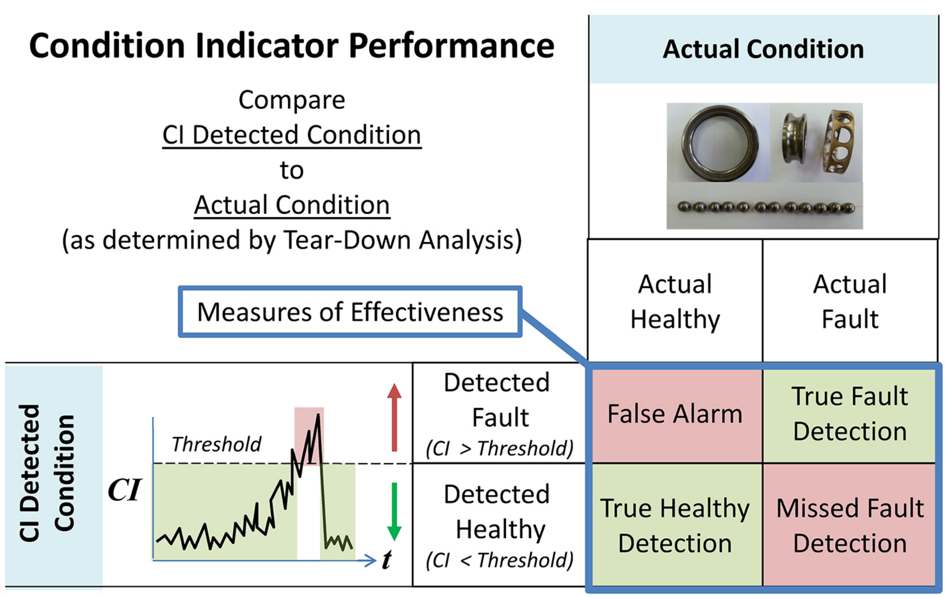


Figure 3. Measuring CI Performance through Comparison against Truth Data

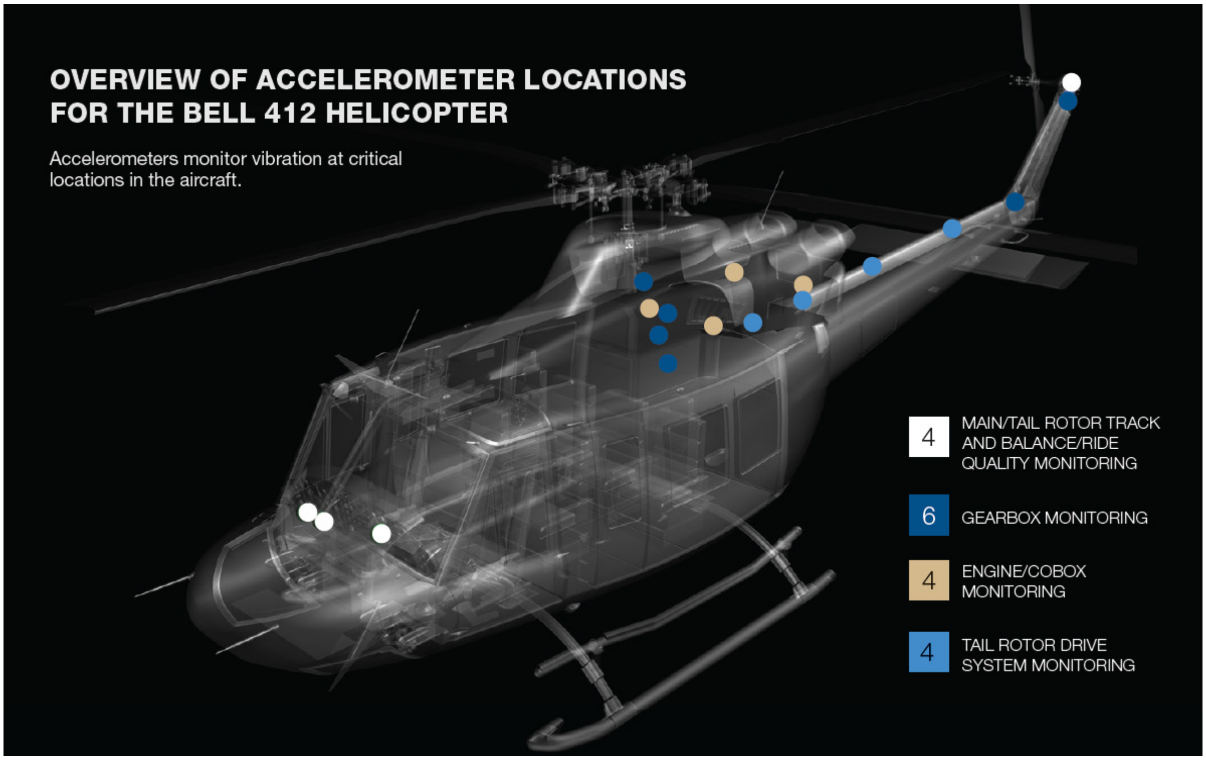


Figure 4. Bell 412 HUMS Sensor Locations

As a secondary objective, this exercise uses the CI/TDA assessment to adjust CI thresholds to improve fault detection performance. Although CI thresholds are established for the 412EP HUMS aircraft, the TDA data provide the opportunity to determine whether those CI thresholds are reasonable. An optimal CI threshold value will provide the operator with sufficient time to perform a maintenance action, but at the same time, avoid false alarms that would call for premature maintenance.

Data Sample

With over 150 HUMS equipped Bell 412EP helicopters in service (some for over 10 years), only 12 aircraft encountered a HUMS detected fault in their gear boxes. To conduct the evaluation of CIs these 12 fleet aircraft were assessed along with 8 fleet aircraft that had no detected faults. The population of 20 HUMS equipped 412EPs used in this assessment is illustrated in Figure 5. The eight aircraft shown in green had no detected faults and served as "control specimens." These included aircraft with known good condition from TDAs as well as some randomly selected aircraft from the population of all aircraft with no known faults. CIs in 12 other aircraft (shown in red) indicated a HUMS-detected fault that was confirmed by TDA of the transmission or gearbox. The Tear Down Assessments confirmed the state of wear/damage on internal components. For the 12 aircraft with faults, CIs detected the faults and triggered gearbox inspections and removals. In one case, a gearbox chip indication was also present at the same time as the CI detection.

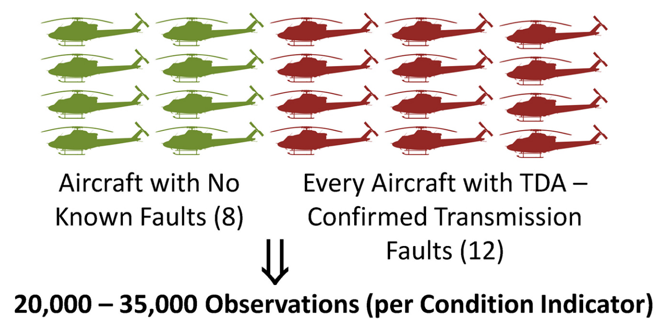


Figure 5. From the fleet of over 150 HUMS-Equipped Bell 412EP helicopters, only 12 aircraft encountered a gear box fault in the course of 10-years of fleet data collection. The Sample Population analyzed these 12 aircraft and 8 others.

For the 20 selected fleet aircraft, the HUMS data generated 20,000 to 35,000 observations per CI over the years of aircraft flying. Since not every CI is calculated for every vibration data sample, the number of observations varies between CIs.

Reclassifying Data based on TDA

HUMS use thresholds to classify CI data as healthy or faulted. To judge the effectiveness of a threshold, a comparison against an independent standard (truth data) is necessary. In this study, the TDA provides insight into the actual condition of the component. To assess the times during which a transmission is "actually healthy" or "actually faulted," a timeline is established in Figure 6 that

represents the gearbox fleet history. Here, we identify three points in the life of the gearbox. Starting with the upward trend in the CI at some earlier time (recorded by the HUMS), a TDA occurs in the middle of Figure 6 soon after the CI threshold is exceeded. As denoted in Figure 6, the gearbox is removed and repaired/replaced. The TDA date and gearbox replacement date are known from the aircraft maintenance records.

Using the three identified points in time, the actual condition of the transmission is assessed as shown in Figure 7. The time before the CI trend change (left hand side of Figure 7) is categorized as healthy. Also, the time after the transmission is repaired or replaced (right hand side of Figure 7) is categorized as healthy. During the intermediate time between CI trend change and gearbox removal (center of Figure 7), the component is in transition from healthy to faulted states.

Since many mechanical faults develop over a period of time, a binary classification of "healthy" or "faulted" is an oversimplification that does not recognize the continuous nature of fault progression. A reasonable estimate of the "fault initiation" (the point of transition from healthy to faulted) is necessary to quantify the CI performance. The criterion that establishes "fault initiation" can significantly impact the results of the analysis that follows. Engineering judgment led to the classification method shown in Figure 8. For the period of time identified as "transition," a number of CI observations (N) are made (as indicated in the lower center of Figure 8). Since HUMS observations are taken on a schedule during certain flight regimes, the number of observations is proportional to the aircraft operating time. Some of the time in the transition period is categorized as "healthy" (the first 2/3 of the N observations) and some as "faulted" (the last 1/3 of the N observations). This "2/3-1/3" method was applied in cases where the CI data exhibited an increasing trend over time (rather than a sudden step change).

This method of classification is applied to each of the 20 aircraft's CI data. For the eight aircraft with no known faults, all CI data were assumed to be healthy as shown to the left of Figure 9. For 12 aircraft with TDA-confirmed faults (right side of the Figure 9), the classification of the actual condition allowed the CI data to be separated into both healthy and faulted populations. In this manner, CI data from all study aircraft were accumulated into separate healthy populations and faulted populations.

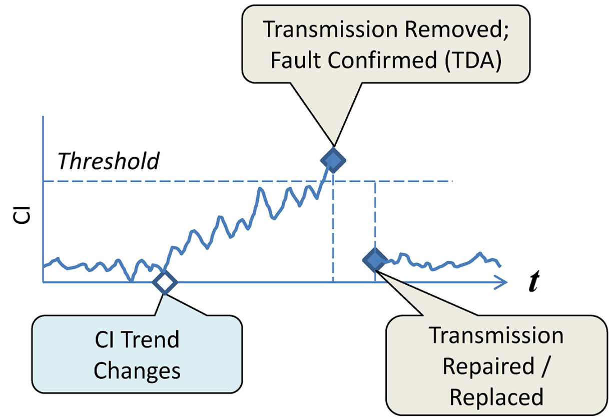


Figure 6. Event Timeline and CI Data Time History

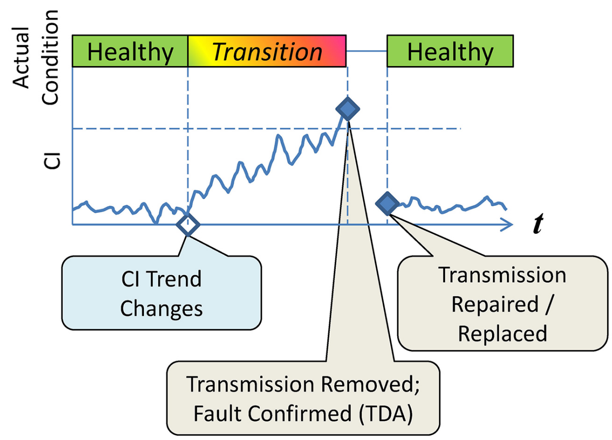


Figure 7. Initial Classification of Actual Condition of Transmission Based on Known Events

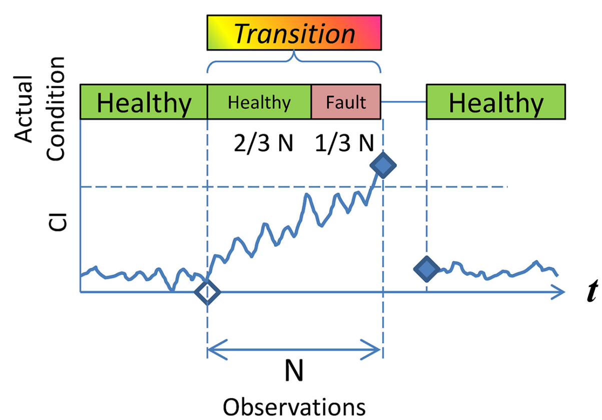


Figure 8. Final Classification of Actual Condition of Transmission

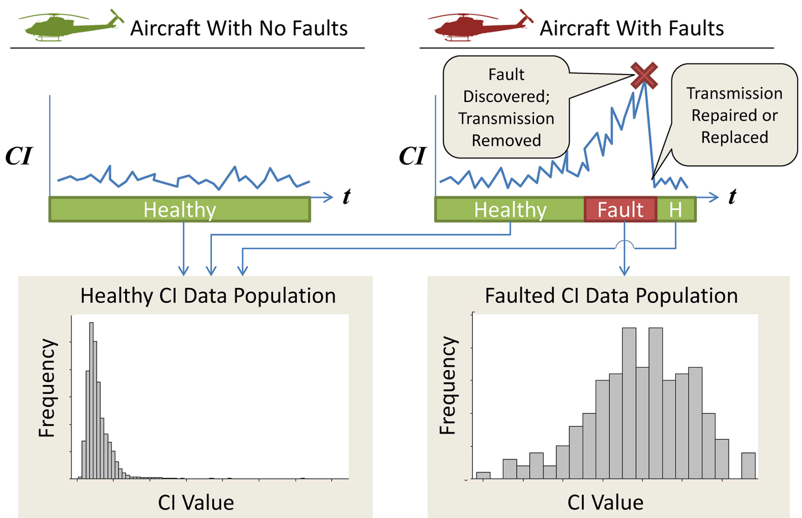


Figure 9. CI Data Separation into Healthy and Faulted Populations. Only faulted data are accumulated in the "Faulted CI" populations.

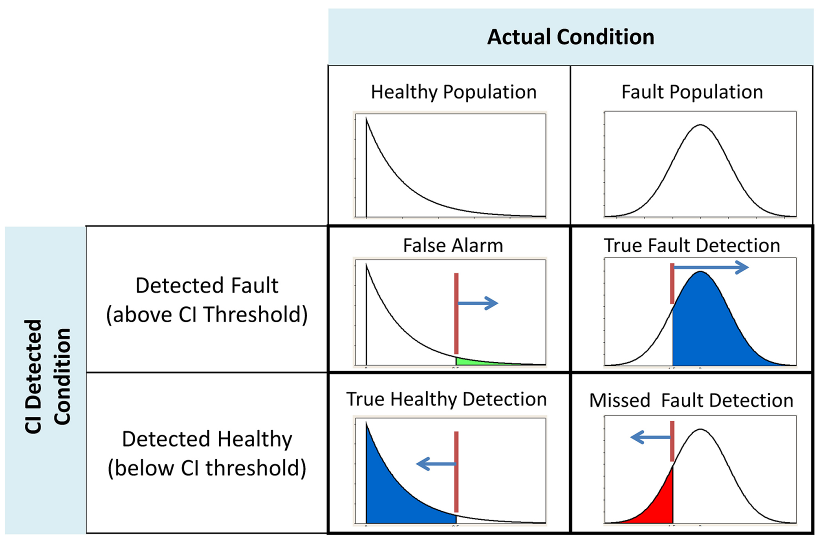


Figure 10. Analysis of Healthy and Fault Populations to Determine CI Performance

Further Separation and Analysis

The next step of the analysis quantifies the CI measures of effectiveness. The general approach used to determine these measures is shown in Figure 10 (which applies the concepts from Figure 3). Starting from the upper portion of Figure 10, the healthy and faulted populations of CI data (left and right columns respectively) are each separated into two groups based on the detected condition (on the left hand side of Figure 10). In the middle row, CI data that are above the CI threshold (detected fault) in each population are identified. CI data that are below the CI threshold (detected healthy) are in the bottom row. For the healthy population, any CI data that falls above the threshold is a false alarm (upper left quadrant of the matrix). For the faulted population, any CI data that falls below the threshold is a missed fault detection (lower right quadrant). The other two cases (upper right and lower left quadrants) represent correct classification.

The specific approach used to determine the measure of CI effectiveness is shown in Figure 11. First, a distribution fit is made for each of the CI data populations as shown in the top of Figure 11. Moving down Figure 11, the distribution properties are used to determine the probability of a missed detection and the probability of a false alarm based on the current (normalized) CI threshold of 1.0.

To evaluate the impact of changing a CI threshold, a receiver-operator characteristic (ROC) curve [11] was developed for some of the helicopter components. Statistical distributions of the healthy and faulted CI data are used to generate ROC curves. ROC curves represent a "sweep" through a wide range of potential threshold values and show missed detection and false alarm probabilities at each value. Youden's J statistic [15, 16] was used to determine the optimum threshold value to maximize the true positive rate and minimize the false positive rate. This statistic will be discussed further in the Application section of the paper.

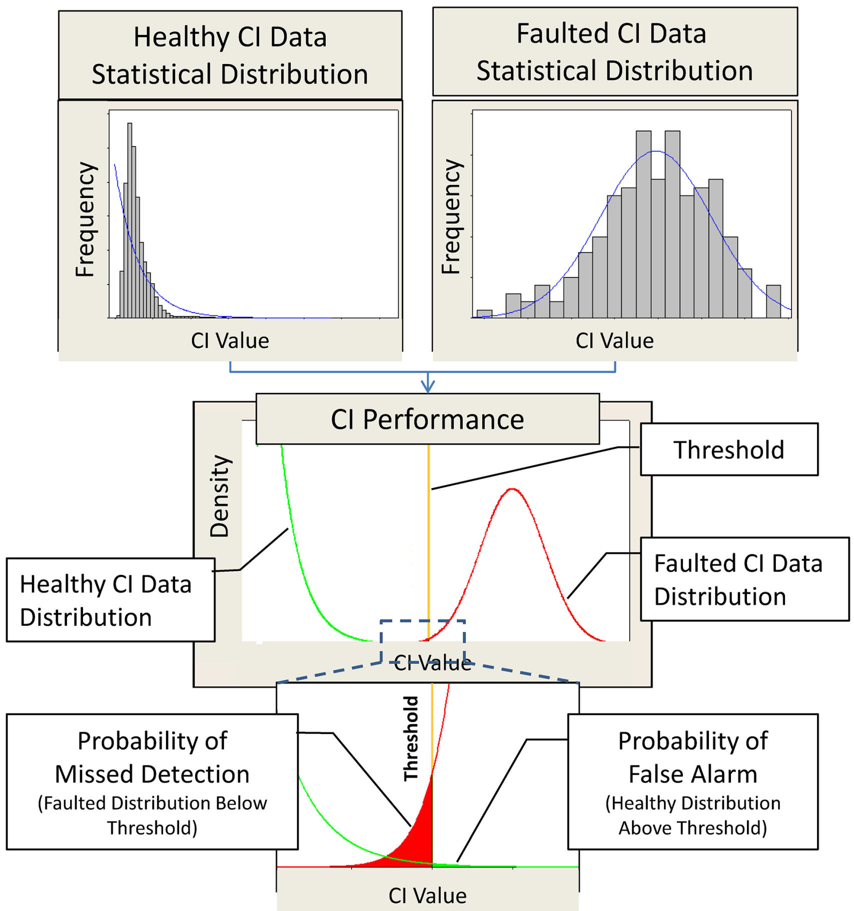


Figure 11. Analysis of Healthy and Faulted Populations to Determine CI Performance

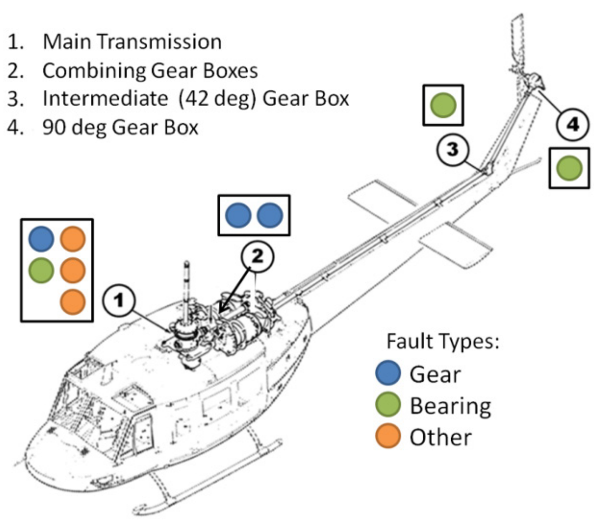


Figure 12. Bell 412 Transmission and Gearbox Locations and Faults Found in Sample Population

RESULTS

The transmission and gearboxes from the 412EP study are identified with their associated fault types in Figure 12. For the 12 aircraft with faults, the TDAs found 9 fault types. Of these 9 fault types, one type (the intermediate gearbox output bearing fault) occurred in four aircraft. This fault is identified as item 3 in Figure 9. All of the remaining 8 fault types were found in separate aircraft. Of the 9 fault types, three were associated with bearings, three were associated with gears, and three were associated with couplings and case structure (labeled "other" in Figure 12). While all gear boxes have metallic "chip" detectors, there was only one instance where an elevated CI was associated with an actual chip detection. This occurred in the 90 degree (tail rotor) gearbox (item 4) which had a chip indication (and elevated CI value) at the time of removal. All other faults were discovered through CI detection alone.

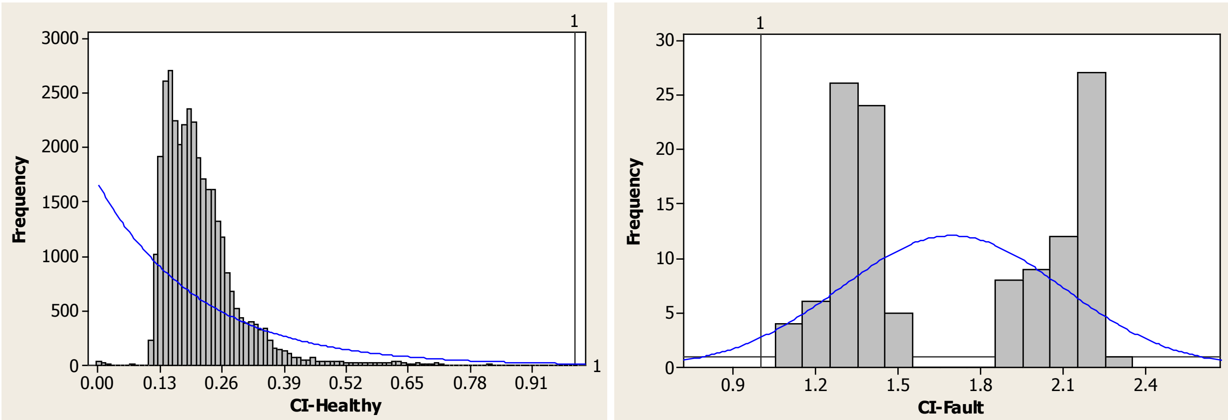
The statistical analysis results for each fault type are listed in Table 2. The data are organized by gearbox and then by faulted component within the gearbox—including the number of occurrences of the fault. The CI algorithm that led to detection of the fault is also listed. For one fault type (the main transmission tail rotor output gear), two CIs are listed since both CIs showed a significant increase at the time prior to component removal. The distribution fits for healthy and faulted data are listed along with the number of CI observations used to establish the fit. Finally, the measures of effectiveness—probability of missed detection and false alarm—are enumerated for the current CI threshold value (1.0).

The process for calculating the CI effectiveness (POMD and POFA) in Table 2 is provided by taking the example of the intermediate gearbox output quill bearing fault (Fault 8 in Table 2). Figure 13(a) and (b) show, respectively, the actual classified healthy and faulted data populations as well as the statistical distributions used to assess the CI performance. These distributions are compared to the CI threshold in Figure 13(c) to illustrate the performance metrics.

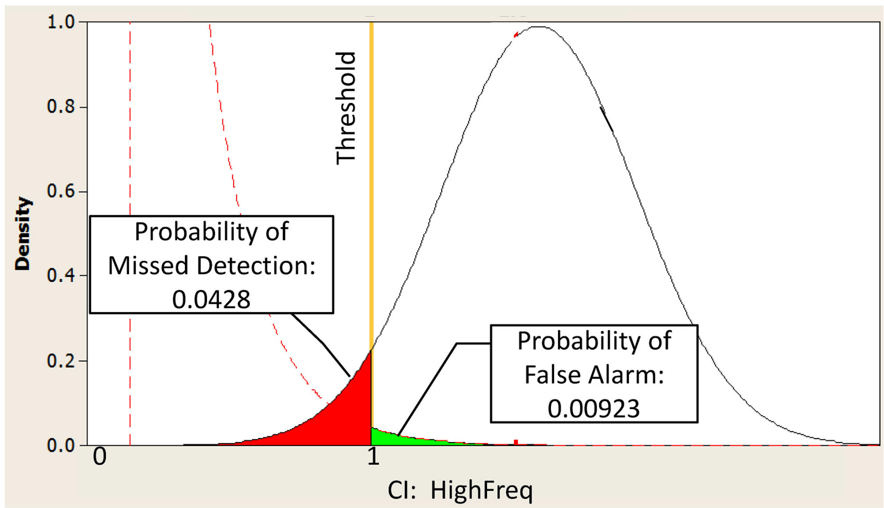
Table 2. CI Performance Summary (for Existing Thresholds).

Fault #	Assembly	Component name (Number of component faults in Study Group)	CI	Distribution Fit				Probability of Missed Detection (POMD)	Probability of False Alarm (POFA)
				Healthy		Fault			
				Type	Data Points	Type	Data Points		
1	Main Transmission	Case Arm (1)	2 X GM	Exponential (2-Parm.)	27,336	Normal	510	50.0%**	1.6%
2	Main Transmission	Tail Rotor Output Quill Bearing (1)	HighFreq	Exponential (2-Parm.)	30,578	Normal	2,053	95.1%**	0.2%
3	Main Transmission	Tail Rotor Output Quill Coupling Adapter (1)	1/rev	Exponential (2-Parm.)	29,230	Normal	106	0.0%*	1.4%
4	Main Transmission	Tail Rotor Output Quill Gear (1)	GM	Exponential (2-Parm.)	29,239	Normal	643	76.2%**	1.0%
			SI	Exponential (2-Parm.)	23781	Normal	643	46.1%**	5.9%
5	Main Transmission	Tail Rotor Output Coupling (1)	GM	Exponential (2-Parm.)	30,121	Normal	99	2.1%	0.6%
6	Engine Combining Gearbox	Intermediate Idler Gear (1)	DA1	Exponential (2-Parm.)	27,366	Normal	480	82.2%**	4.7%
7	Engine Combining Gearbox	Input , Idler Gears (1)	SI	Exponential (2-Parm.)	28,087	Normal	197	0.5%	0.0%*
8	Intermediate (42 degree) Gearbox	Output Quill Bearing (4)	HighFreq	Exponential (2-Parm.)	33,107	Normal	122	4.3%	0.9%
9	Tail Rotor (90 degree) Gearbox	Output Shaft Bearing (1)	High Freq	Exponential (2-Parm.)	34,268	Log-normal	224	93.1%**	0.0%*

* Note: Due to rounding, values less than 0.05% are listed as 0.0%.



(a) Healthy Population Distribution Fit (b) Faulted Population Distribution Fit



(c) Statistical Performance of CI for Current Threshold (1.0)

Figure 13. Intermediate Gearbox Output Bearing – High Frequency CI Characterization

The results of Table 2 were mixed. With the methodology described, the likelihood of a missed detection appears far greater than the likelihood of a false alarm. In fact, all CIs appear to have a very good (low likelihood) false alarm characteristics. For the same methodology, the probability of missed detection was not as good. Four of the fault types were diagnosed with excellent results for POMD. Satisfactory POMD/POFD results are highlighted in green in Table 2. There was no distinct pattern among CI algorithms or type of component which could explain the difference in performance; each algorithm and component type (gear, bearing, other) had both good and bad performers. Even in the cases where the CI performance appears poor, all faults were in-fact discovered by the HUMS CI data. Since the statistical results appear anomalous and contradictory to the fleet HUMS data, additional investigation is necessary.

A second look at the low correlation CIs revealed that each fault developed steadily over a long period of time (in some cases, several months). A slow-moving trend put these faults at a disadvantage in the analysis. Referring back to the CI data categorization from Figure 8, the "2/3-1/3" categorization method is applied to a slow-moving CI trend in Figure 14. The CI threshold is shown by the horizontal orange line. The right side of the CI history shows that the customer removed the transmission shortly after the CI exceeded its threshold. From the point of the CI trend change, Figure 14 shows the application of the previously discussed "2/3-1/3" categorization method. All of the CI data above the red "fault" is categorized as faulted. The distribution of the faulted population data are shown to the right of the CI plot (Figure 14). Since the CI was "slow-moving," a large portion of the observations were classified as faulted even though the observations lie below the CI threshold; this portion of the distribution is highlighted in blue. The portion of the total faulted population that falls in the blue region represents the probability of missed detections. In this case, this probability is much more than 50% (poor performance). Therefore, while the fault in this example was detected using the existing threshold, the categorization method used in this study yielded poor statistical performance. The Lessons Learned section will discuss potential alternative categorization methods which may more accurately assess performance for these slow-moving cases.

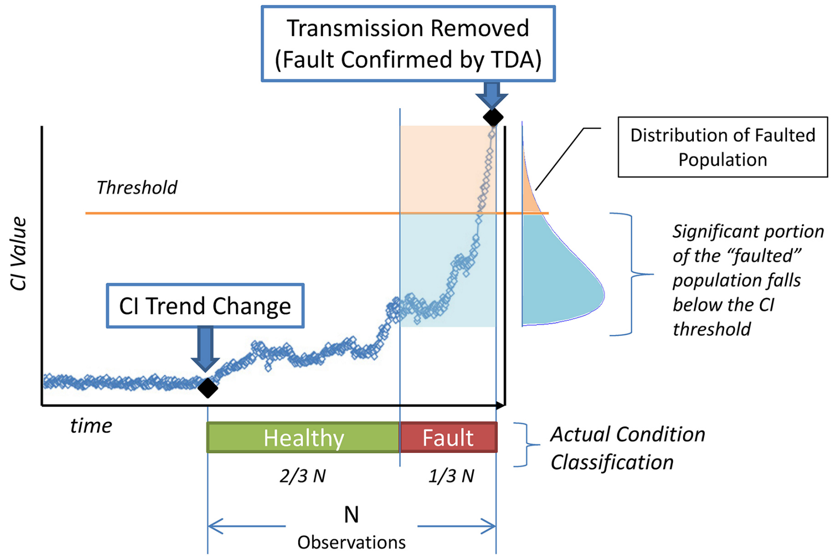


Figure 14. Slow Moving CI Trends Perform Poorly Using Proposed Classification Method

APPLICATION

The goal of CI performance analysis is to balance the likelihood of failing to catch a pending failure against a false detection, while not prematurely removing faulted components. The simplest adjustment is a change in the CI threshold. In the case of a faster-moving CI trend (which showed good performance using the approach in this study), it may be possible to optimize the threshold to improve the probability of detection of faults while maintaining a reasonably low probability of failure. To evaluate changes in threshold, a receiver operator characteristic (ROC) curve for the intermediate gearbox output quill bearing fault is shown in Figure 15. Here, the false positive rate is plotted against the true positive rate (which is the complement to missed detection). A change in threshold affects both true positive and false positive rates. A high threshold will decrease false positives, but it will also reduce true positive rate (and thus increase missed detections). Conversely, a low threshold will maximize the true positive rate (and minimize missed detections), but it will also lead to a higher false positive rate. The dashed line in Figure 15 represents "random chance"—equal false positive and true positive rates (equivalent to a coin toss; the worst case). A perfect classifier would be a point in the upper left corner with 1.0 true positive rate and no false positives. In reality, every CI will have an ROC curve between the two extremes. To quantify performance, Youden [15] proposed the J statistic—a measure of the vertical distance between the ROC curve and the "random chance" line. This measurement can be used to find the point on the ROC curve that is furthest from random chance. This is the point with the highest diagnostic effectiveness.

Figure 16 selects the region of interest from the ROC curve in Figure 15. The highlighted (circled) point on the left of Figure 16 shows the performance for the current normalized CI threshold of 1.0 (which was listed as Fault 8 in Table 2). As shown in Figure 16, the probability of missed

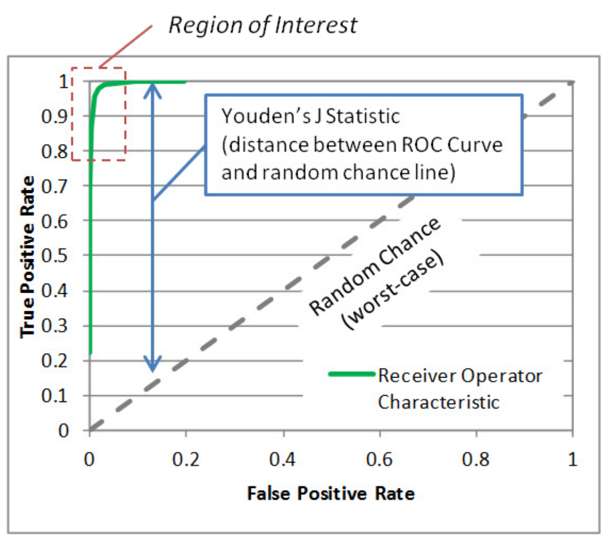


Figure 15. Receiver-Operator Characteristic (ROC) Curve for Intermediate Gearbox Fault (Fault 8 in Table 2)

detection is 4.3% and the probability of False alarm is 0.9% for the current CI threshold. The highlighted (circled) point on the right represents the CI performance at a lowered threshold of 0.85 (instead of 1.0). This point has the highest diagnostic effectiveness based on Youden's J statistic and yields a balanced 1.9% probability of missed detection and the same probability of false alarm (1.9%).

It is tempting to apply the same (reduced threshold) methodology to the other CIs with poor POMA/POFA balance to yield better statistical results. However, doing so may not ultimately improve CI performance. Due to rules in place for commercial operators, changing the CI threshold value will necessarily change the point at which a component is removed. Considering the case previously discussed in Figure 14, it would appear (on the surface) that a simple threshold reduction indicated in Figure 17 by moving from the solid horizontal line down to the dashed

horizontal line, would significantly improve CI performance by moving the majority of the faulted population above the threshold. While doing so would greatly reduce the probability of a missed detection, a commercial operator would likely face earlier maintenance actions and higher DOC. As shown in Figure 18, the likely outcome of a lower CI threshold is simply earlier part removal. The results of applying the lower threshold produce a new classification of healthy/faulted that looks very similar to the baseline case before the threshold was changed. While the amount of data between fault initiation and component removal has been reduced, a similarly large percentage of the "faulted" population is still below the new threshold. This would again lead to a high probability of missed detection—which amounts to poor balance between POFA and POMA.

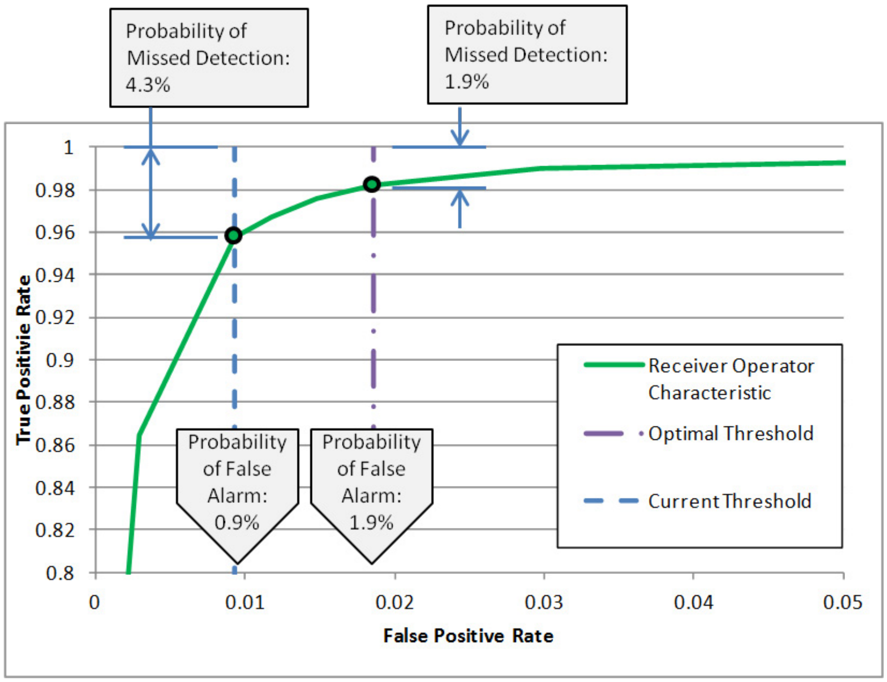


Figure 16. Region of Interest from Figure 15 Showing Significant Improvement in Missed Detections by optimizing the CI fault detection threshold

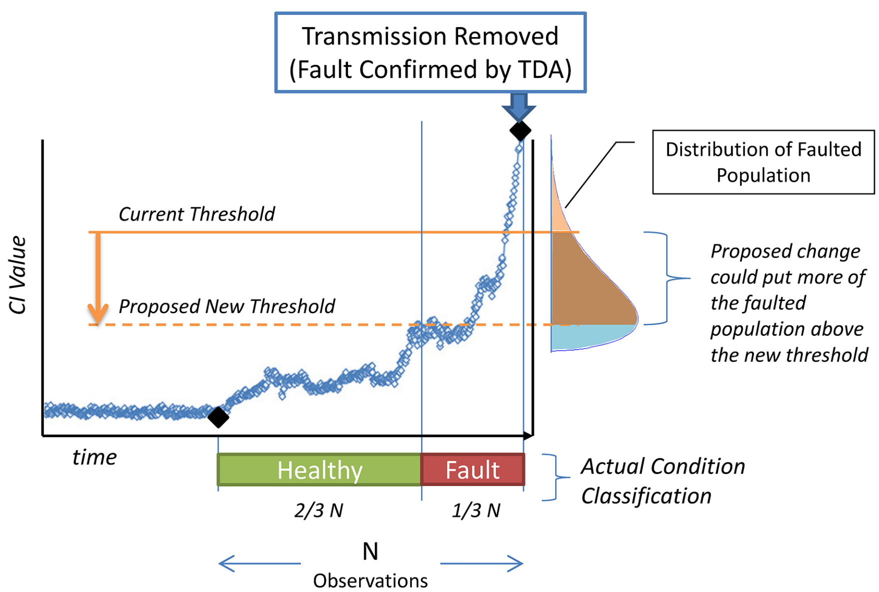


Figure 17. Idealized Performance Improvement due to Threshold Change

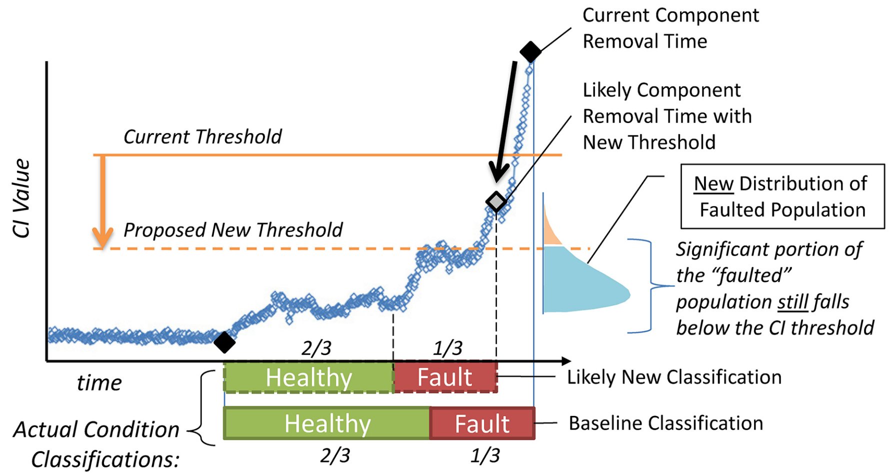


Figure 18. Actual Threshold Change Impact Yielding Little Change in Performance

LESSONS LEARNED

The primary lesson learned in this study is that the classification method (2/3-1/3 healthy/faulty ratio identifier and CI threshold setting) for creating the healthy and faulted populations has mixed results. While all of the CIs were successful in detecting the faults at the point the components were removed, faults that trended slowly towards their CI threshold performed poorly in the statistical performance of POFA/POMA. This observation leads to the conclusion that a new classification method for taking account of CI changes with time may be necessary to better represent slow growing faults. Time-dependent receiver operator characteristic analysis [17] may be a means to achieve this end and is recommended for future work.

Not all of the lessons learned in this study were technical. The process of gathering the data for this study also yielded several lessons learned for data-sharing and communication between the helicopter OEM and helicopter operators. In all of the faults considered in this study, the tear-down assessments were performed either by the aircraft operator or by a third-party MRO provider – not by the OEM. Most operators understand the value of uploading their HUMS data to the OEM so that Product Support Engineers (PSE) can review the data and provide recommendations. The value of providing detailed tear-down assessments to the OEM is less apparent to the customer. Yet, the results have longer-term benefits to the customer because the TDA information can be used to establish improved HUMS indications and better component replacement strategies. Additional education for the operator community as to the value of communicating the actual part condition is recommended to promote HUMS improvement.

CONCLUSIONS

A statistical analysis of fleet data was used to quantify diagnostic effectiveness of HUMS condition indicators. In the case of the Bell 412EP study group, several fault types and their associated CIs had good diagnostic performance and the data could be used to determine CI threshold changes that may improve the statistical ability to avoid false alarms or missed fault detections. For slow-growth fault types, the selected improvement methodology did not lead to improved diagnostic performance due to the classification method. Additional work in a methodology that recognizes slow growth faults is needed to better represent the performance of such CIs. Further outreach to HUMS equipped aircraft operators must be emphasized. The importance of sharing operator maintenance data (including tear-down assessments) with the OEM is strongly recommended to improve future HUMS data analysis.

Author contact: Brian Tucker btucker4@bh.com,
Paula Dempsey paula.j.dempsey@nasa.gov

ACKNOWLEDGMENTS

This effort was jointly funded by NASA and the FAA under contract NNC10BA04B. The authors wish to acknowledge the work of Ranga Maddali (Bell Helicopter, retired) in this research and the support of Joe Wendelsdorf (Bell Helicopter, retired).

REFERENCES

[1]McKenna, J. T., "An Eye on Operations," *Rotor & Wing*, September 2005.

[2]Dempsey, P. and Zakrajsek, J., "Rotorcraft Health Management," *System Health Management: with Aerospace Applications,* eds. Johnson, S.B., Gormley, T., Mott, C., Kessler, S., and Patterson-Hine, A. John Wiley & Sons, Chichester, NY, 2011, pp. 577-585.

[3]Stewart, R.M. "Some useful data analysis techniques for gearbox diagnostics," Machine Health Monitoring Group,

Institute of Sound and Vibration Research, University of Southampton, Report MHM/R/10/77, July 1977.

[4]Zakrajsek, J.J. “An investigation of gear mesh failure prediction techniques,” NASA TM–102340, AVSCOM TM 89–C-005, 1989.

[5]Crawford, A.R. and Crawford, S., *The Simplified Handbook of Vibration Analysis*, Computational Systems Incorporated, 1992.

[6]Antolick, L., Branning, J., Dempsey, P. and Wade, D., “Evaluation of Gear Condition Indicator Performance on Rotorcraft Fleet,” American Helicopter Society International 66th Annual Forum, Phoenix, AZ, 11-13 May 2010.

[7]Howard, Ian, “A Review of Rolling Element Bearing Vibration “Detection Diagnosis and Prognosis,” DSTO Aeronautical and Maritime Research Laboratory, DSTO-RR-0013, October 1994.

[8]Delgado, I., Dempsey, P., Antolick, L. and Wade, D., “Continued Evaluation of Gear Condition Indicator Performance on Rotorcraft Fleet,” American Helicopter Society Airworthiness, CBM, and HUMS Specialists' Meeting, Huntsville, AL, February 11-13, 2013.

[9]Dempsey, Paula J., Wade, Daniel R., Antolick, Lance J., and Thomas, Josiah, “Investigation of Spiral Bevel Gear Condition Indicator Validation via AC–29–2C Using Fielded Rotorcraft HUMS Data,” NASA/TM—2014-218406, November 2014.

[10]Allen, J., “CBM Vibration Monitoring- Lessons Learned from the Apache MSPU Program”. Presented at the American Helicopter Society Airworthiness, CBM and HUMS Specialists’ Meeting, Huntsville, AL, February 9-11, 2015.

[11]Dempsey, P., Keller, J. and Wade, R., “Signal Detection Theory Applied to Helicopter Transmission Diagnostic Thresholds,” NASA/TM—2008-215262; AMRDEC PAO Control Number FN 3597, July 2008.

[12]UK Civil Aviation Authority CAP 753, “Helicopter Vibration Monitoring (VHM) - Guidance Material for Operators Utilising VHM in Rotor and Rotor Drive Systems of Helicopters,” Version 1, August 2012.

[13]Helioffshore, “Helioffshore HUMS Best Practice Guidance,” Version 1.0, November 2015.

[14]Wendelsdorf, J., “Development And Fielding Of Affordable HUMS Technology For Commercial Helicopters,” Presented at Heli Japan 2010, Saitama, Japan, November 1-3, 2010.

[15]Youden, W.J., “Index for rating diagnostic tests,” *Cancer* Vol. 3, 1950, pp.32-35.

[16]Schisterman, E. F., Perkins, N. J., Liu, A., & Bondell, H., “Optimal cut-point and its corresponding Youden index to discriminate individuals using pooled blood samples. *Epidemiology*, Vol. 16(1), 2005, pp.73-81.

[17]Heagerty, Patrick J., Thomas Lumley, and Margaret S. Pepe. “Time-dependent ROC Curves for Censored Survival Data and a Diagnostic Marker”. *Biometrics* Vol. 56.2, 2000, pp. 337–344.